R E P O R T  R E S U M E S

ED 012 154 FL 000 321

TESTS WITH A NEW LOOK AND A NEW PURPOSE--THE MLA COOPERATIVE FOREIGN LANGUAGE TESTS.

BY- BRYAN, MIRIAM M.
NATIONAL EDUCATION ASSN., WASHINGTON, D.C.

PUB DATE DEC 66

EDRS PRICE MF-$0.09 HC-$0.28 7P.

DESCRIPTORS- *LANGUAGE TESTS, *LANGUAGE PROFICIENCY, *STANDARDIZED TESTS, *MODERN LANGUAGES, *TEST CONSTRUCTION, READING SKILLS, WRITING SKILLS, AUDIOLINGUAL SKILLS, DISTRICT OF COLUMBIA, MODERN LANGUAGE ASSOCIATION, EDUCATIONAL TESTING SERVICE, COOPERATIVE FOREIGN LANGUAGE TESTS, NDEA TITLE VI, PRINCETON, MLA

UNDER TITLE VI OF THE NATIONAL DEFENSE EDUCATION ACT THE MODERN LANGUAGE ASSOCIATION OF AMERICA (MLA) UNDERTOOK IN 1960 TO PREPARE A SERIES OF TESTS IN FRENCH, GERMAN, ITALIAN, RUSSIAN. AND SPANISH FOR USE OF TWO LEVELS--(1) AFTER THE 2D YEAR OF HIGH SCHOOL LANGUAGE STUDY, OR THE SECOND SEMESTER IN COLLEGE, AND (2) AFTER THE 4TH YEAR OF HIGH SCHOOL LANGUAGE STUDY OR THE FOURTH SEMESTER IN COLLEGE. SUPERVISION OF THE PROJECT WAS ENTRUSTED TO OUTSTANDING FOREIGN LANGUAGE TEACHERS, WHO WORKED INITIALLY IN COOPERATION WITH THE EDUCATIONAL TESTING SERVICE (ETS) IN PRINCETON. THE 2D COMMITTEES FORMED PRODUCED A BATTERY OF TESTS, PRE-PRETESTED AS WELL AS PRETESTED, WHICH MEASURE LISTENING, SPEAKING, READING, AND WRITING PROFICIENCY IN EACH OF THE FIVE LANGUAGES. ALTHOUGH THE SCORING SERVICES OF THE MLA-ETS ARE AVAILABLE, DETAILED INSTRUCTIONS FOR SCORING ACCOMPANY EACH TEST IN ORDER TO ELIMINATE SUCH EXPENSE. THE RESULTS SUGGEST THAT--(1) THE TESTS ARE OF SUITABLY GREATER THAN MIDDLE DIFFICULTY, (2) THE PROGRESSION FROM LEVEL TO LEVEL IS WELL PLANNED, (3) STUDENTS IN TRADITIONAL COURSES SHOW SOME, BUT NOT MARKED, SUPERIORITY IN READING AND WRITING, AND (4) WHILE STUDENTS WITH 2 YEARS OF A LANGUAGE IN HIGH SCHOOL DO ABOUT AS WELL AS THOSE WITH 1 YEAR OF COLLEGE, THOSE WITH 4 YEARS IN HIGH SCHOOL DO SLIGHTLY BETTER THAN THOSE WITH 2 YEARS IN COLLEGE. THIS ARTICLE WAS PUBLISHED IN "THE DFL BULLETIN," VOLUME 6, NUMBER 2, DECEMBER 1960. (GJ)

ED012154

BULLETIN BULLETIN BULLETIN BULLETIN

# the DFL BULLETIN

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

FL 000 381

## Featured in This Issue:

*The Role of the Cooperating Teacher*
*NDEA Foreign Language Leadership Institute*
*Tests with a New Look*
*Let's Put the Accent on Success*
*A Boost for the Classics*

**DFL BOARD OF DIRECTORS**

**GEORGE E. SMITH, President,** Indiana Language Program, Bloomington.

**LESTER W. McKIM, President-Elect,** Central Washington State College, Ellensburg.

**GENEVIEVE S. BLEW, Past President,** Maryland State Department of Education, Baltimore.

**FRANK M. GRITTNER, Secretary,** Wisconsin State Department of Public Instruction.

**ELIZABETH H. RATTÉ, Vice-President** (Elementary), Purdue University, Lafayette, Indiana.

**MARIANNE C. CIOTTI, Vice-President** (Junior High School), Vermont State Department of Education, Montpelier.

**ANDREA McHENRY MILDENBERGER, Vice-President** (Senior High School), City University of New York.

**NELSON H. BROOKS, Vice-President** (College), Yale University, New Haven, Connecticut.

**REGIONAL REPRESENTATIVES**

**REMUNDA CADOUX, Region II** (N.J., N.Y., Pa.): French Dept., Christopher Columbus High School, Bronx, New York.

**LUCILLE COX, Region III** (Del., Md., Va., W. Va., N.C., D.C.): Classical Languages, E. C. Glass High School, Lynchburg, Va.

**ELIZABETH F. BOONE, Region IV** (S.C., Ga., Ala., Miss., Fla., La., Ky., Tenn., Puerto Rico): Supervisor of Secondary Foreign Language, Dade County Public Schools, Miami, Fla.

**VIOLET E. BERGQUIST, Region V** (Ohio, Ind., Ill., Mich., Wis., Minn., Iowa, N.D., S.D.): Education Dept., University of Illinois, Chicago.

**HELEN SHELTON, Region VI** (Wash., Ore., Idaho, Mont., Wyo., Utah, Alaska): Supervisor Foreign Language Program, State Department of Public Instruction, Olympia, Washington.

**MARIE ESMAN BARKER, Region VII** (Kans., Mo., Okla., Colo., Neb., Ark., Tex., N. Mex.): Special Project Director for Title III, El Paso (Tex.) Public Schools.

**MARY I. MONTGOMERY, Region VIII** (Cal., Nev., Ariz., Hawaii): Teacher of Spanish, South Gate Junior High School, Los Angeles, Calif.

**EDITORIAL AND BUSINESS STAFF**

**JERMAINE D. ARENDT, Co-Editor,** Minneapolis Public Schools, Minneapolis, Minnesota.

**EMMA MARIE BIRKMAIER, Co-Editor,** University High School, University of Minnesota, Minneapolis, Minnesota.

**ELIZABETH RATTÉ, FLES Editor,** Purdue University, Lafayette, Indiana.

**CLEMENS HALLMAN, Managing Editor,** Indiana University, Bloomington, Indiana.

BULLETIN the DFL BULLETIN

Published quarterly by the Department of Foreign Languages of the National Education Association, 1201 Sixteenth St., N.W., Washington, D.C. 20036.

*The DFL Bulletin solicits and publishes original articles relating to the teaching of foreign languages. Points of view expressed in these articles may support or oppose positions taken by the Department of Foreign Languages as an organization.*

**CONTENTS OF THIS ISSUE** **Page**

1. The Role of the Cooperating Teacher ..... 2

2. NDEA Foreign Language Leadership Institute ..... 4

3. Tests with a New Look and a New Purpose ..... 6

4. Let's Put the Accent on Success! ..... 9

5. A Boost for the Classics ..... 10

VOLUME VI • NUMBER 2 • DECEMBER 1966

Entered as third class matter at the post office in Washington, D.C. Subscription rate, $3.00 per year. Single copies, when available, 50¢ each. Order from the Department of Foreign Languages, 1201 Sixteenth Street, N.W., Washington, D.C. 20036.

Address all inquiries regarding articles to the editors. Questions regarding advertising should be directed to the Managing Editor.

## THE MLA COOPERATIVE FOREIGN LANGUAGE TESTS:

# Tests with a New Look and a New Purpose

*by Miriam M. Bryan*
Educational Testing Service
Princeton, New Jersey

As recently as a decade and a half ago, most of our foreign language instruction in both schools and colleges was aimed at the development of competence in reading, and the standardized tests used to measure language proficiency were essentially reading tests. Today, although competence in reading is still an important objective, our foreign language instruction is aimed at the development of competence in all four language skills: listening, speaking, reading, and writing. Just as the emphasis on the development of complete language competence has led to the development of new instructional procedures and new learning materials, so, inevitably, it has led to the development of new types of testing materials to measure the products of the new approach. The story of the development of these new testing materials is essentially the story of the development of the MLA Cooperative Foreign Language Tests.

In 1960 the Modern Language Association of America received a contract from the United States Office of Education under Title VI of the National Defense Education Act for the development of a series of tests suitable for the evaluation of language learning by the audio-lingual approach. The amount of the contract was more than half a million dollars.

Major specifications for the tests were as follows:

1. Tests should be developed in the five languages most frequently taught in American secondary schools and colleges — French, German, Italian, Russian, and Spanish.
2. Tests should measure the four language skills—listening, speaking, reading, and writing.
3. Tests should measure skills on two levels, the lower level corresponding to the first and second year of language learning in the secondary school or to the first and second semester of language learning in college, and the higher level corresponding to the third and fourth year of language learning in the secondary school or to the third and fourth semester of language learning in college.
4. There should be two forms of each test at each level.

To assure that the tests would demonstrate the best practices in language teaching, the task of planning and writing them was assigned by the MLA to committees of outstanding foreign language teachers. Professor Nelson Brooks of Yale University was appointed MLA director of the testing program. Education Testing Service was invited to furnish the testing "know-how." Mr. Donald D. Walsh, director of the Foreign Language Program for the MLA, assumed overall responsibility for the program.

So as to obtain the broadest possible representation, the committee members were drawn from all parts of the country, from secondary schools and colleges, from public and private institutions. Twenty committees were formed, one for each skill in each language. Each committee consisted of three members, preferably one college teacher, one public school teacher, and one private school teacher. On most committees at least one member was a native speaker of the language to be tested.

It is important for foreign language teachers to know about these committees because it is important for them to be assured that these tests were not developed by "testers" and not by novices at foreign language teaching, but by outstanding members of the foreign language teaching profession under the direction of the highest ranking professional organization in the field. ETS foreign language specialists served as consultants to the test committees through the test planning and item writing stages, assisted the test committees in the interpretation of item analyses and in the assembly of test forms, and planned and conducted the pretesting and standardization programs.

The chairmen of the twenty committees formed an advisory council, which met in the fall of 1960, to discuss and draw up the broad outlines of the program. The committees then met in skill groups to put together test specifications for each skill separately, to draw up content outlines for the tests and to begin the actual test construction. They concluded the major portion of their work three years later, in the spring of 1963, when final forms of the tests were approved for standardization.

The tests are designed to measure competence in all four language skills in a functional context. In the listening tests, the students listen; in the speaking tests, they speak; in the reading tests, they read; and in the writing tests, they write. Except for directions, which are given in English, the tests are entirely in the language tested. They contain only complete and natural utterances. Isolated vocabulary, artificial phrases, and wrong forms are avoided. All taped material is spoken by native voices. The tests in all four skills present a wide variety of contexts, with gradual progression from the very simple to the very difficult. While the tests were designed to fill the need for evaluative instruments in schools using the audio-lingual approach, they can be used in any school that has the equipment needed to administer them. From the most traditional school, in which only the reading tests or the reading and

writing tests may be appropriate to the most progressive school, where tests in all four skills are appropriate, the tests are useful in varying degrees.

**LISTENING:** In the listening test, which is received through earphones, the student is required to answer multiple-choice questions based on single utterances, on short conversations between speakers, on passages of connected discourse read by a single speaker, on telephone conversations where the examinee assumes the part of one of the speakers, and on brief dramatic scenes enacted by several voices, men's and women's. In the lower level tests, the first few questions involve visual stimuli and responses. The student may record his answers in the test booklet or on a separate answer sheet.

**SPEAKING:** The speaking test allows for student reception from master tapes through individual recording stations; student responses are on tape. The student is tested first on ability to repeat what is heard with the proper pronunciation and intonation, then on ability to read aloud with the proper pronunciation and intonation and with the fluency expected of the student at his particular level of language study, then on ability to give single sentence responses to questions based on picture stimuli, and, finally, on ability to describe pictures presented singly and in sequences. The pictures in the speaking tests are the same for all five languages.

**READING:** The reading test is a pencil and paper test of the multiple-choice type. The first part of the test is aimed at testing knowledge of high frequency words and idiomatic expressions; in this part the questions are presented in sentence completion form. The second part includes reading passages of varying length with multiple-choice questions which test word or phrase discrimination getting the main idea, finding details, and drawing conclusions. Materials for both sentences and passages are from newspapers, periodicals, and literary works within the ability of students to comprehend. The student may record his answers in the test booklet or on a separate answer sheet.

**WRITING:** The writing test is a pencil and paper test of the subjective type. The first part of the test is aimed at testing such elements as articles, prepositions, pronouns, and auxiliaries at the sub-sentence level with items of the fill-in type. In the next part the student is asked to rewrite sentences, making changes of tense, gender, number, person, word order, and sentence structure, the type of change being indicated by clue or example. Finally, the student writes short dialogues or structured paragraphs based on verbal stimuli.

The listening and reading tests are objectively scored, whether the answers are recorded in the test booklet or on a separate answer sheet. The speaking and writing tests require subjective scoring.

The test development program involved so many departures from traditional ways of testing language achievement that an extensive pretesting program and a most comprehensive standardization program were necessary. As a preliminary step, in the spring of 1961, one test in each skill in each language was prepretested in a small number of schools, all using the audio-lingual approach, so that the committees could see if the tests were working in the right direction. When the results of this prepretesting were analyzed, the committees redesigned test specifications as they felt necessary and proceeded to develop three forms of each test for each level, a total of 120 tests.

These tests were pretested in the spring of 1961 in 100 public and private secondary schools possessing the facilities for administering all tests. Approximately 40,000 tests were administered to 10,000 students, each student taking tests in all four skills so that the relative difficulty of the tests from skill to skill could be determined. A detailed item analysis was based on the results of the pretesting. After examining the item analysis data, the committees discarded the one third of the items which proved to be least effective and assembled 80 final forms, in which only items of the highest statistical validity and of the proper difficulty were retained.

Approximately 80,000 final forms were administered in the spring of 1963 in a norming and equating program involving over 20,000 students in more than 400 secondary schools and over 100 colleges. The schools and colleges participating in the standardization program were randomly selected for participation after having previously supplied exhaustive information about their language programs which made it possible to classify them as audio-lingual or traditional — proportion of class time in which the target language was used, amount of translation from the target language into English and vice versa, proportion of time used for the explanation of grammatical constructions in English, and the like.

As a result of this large-scale testing, it has been possible to provide for the lower level tests norms for tests in all four skills separately for high school students who have been taught by the audio-lingual approach, and for tests in reading and writing norms for high school students who have followed a traditional program of instruction. For the higher level forms, only general high school norms are provided, the reason for this being that at the upper grade levels the results of differences in instructional approach are not so clearly pronounced. For the same reason, all college norms are general norms. Norms are available for high school classes at all levels of instruction except for fourth year Italian and Russian at the high school level, and for the speaking test in German at this same level. In spite of a thorough combing of the country for cases for these languages, working through the MLA and through state supervisors of foreign languages, it was not possible to locate enough fourth year cases to permit the production of reliable data for this level. With a single exception, norms are available for second and fourth semester classes at the college level for all tests in all languages; the exception is the speaking test in Italian, for which it was not possible to locate enough cases to permit the development of norms.

The equating resulting from the standardization was both horizontal and vertical, i.e., Form A was equated to Form B at each level and the lower level tests to the higher level tests. As a result of the equating, a score on a test at one level can be converted to a score on the same test at the other level. The ability to perform such a conversion is essential when scores on different test forms at the same or at different levels are being compared.

The most unusual tests are, of course, the speaking tests. The speaking tests represent the first attempt to develop standardized tests of this skill for use with secondary school students and in first and second year college classes. As might be expected, they follow closely the plan for the speaking tests developed earlier for teachers in the series of MLA Foreign Language Proficiency Tests for Teachers and Advanced Students. The test

**TESTS (Cont.)**

tapes for the speaking tests are in two parts. In the first part the test is presented to the students. In the second part an actual student recording is presented to help the teacher in the rating of the tapes.

All the student tapes recorded in the pretesting and standardization programs were scored in the MLA-ETS scoring center by specially trained professional scorers, who were able to achieve a moderately high degree of inter-scorer reliability with recorded responses. While professional scoring services are available in the MLA-ETS scoring center for tapes recorded in local or institutional testing programs, they are not widely used because of their costliness. Teachers who score their own tests completely so that they can compare the test results of their students with those of the norms groups can do a maximum of four tapes per hour. Or teachers may prefer simply to spot-check the student tapes and give a broad grade like A, B, C, D, or Excellent, Good, Fair, Poor. Some teachers have scored their tests either completely or by spot-checking and then sent a small number of tapes to the MLA-ETS scoring center for re-scoring as a check on their own scoring.

The writing tests administered in the pretesting and standardization programs were also rated by the professional scorers in the scoring center. Because of the difficulty of scoring any kind of writing test with any degree of reliability, very detailed directions for the scoring of these tests were worked out for the scorers. Since the tests are highly structured in spite of the fact that they are written tests — and the writing is therefore rated more for correctness than for creativity, the professional scorers were able to achieve an inter-scorer reliability for the writing tests almost as high as that expected for an objective test. While scoring services for the writing tests are available in the MLA-ETS scoring center, the cost is again so high that most scoring is being done locally. This can be quite easily done because of the detailed directions for scoring that are provided — including sample student compositions rated by professional scorers and accompanied by detailed descriptions of the reasons for the ratings assigned. Teachers who score their own writing tests and compare the results of their students with those of the norms groups find that they are able to rate from four to six tests an hour.

Here are a few impressions of the tests and their characteristics, based originally on the results of the standardization program and reinforced by test results reported more recently by test users.

1. The tests are of greater than middle difficulty. This, it should be said, was intended. With the remarkable progress being made today in the teaching of foreign languages at the secondary school level and with the extension of foreign language instruction to the elementary grades, the sponsoring agencies and the committees were afraid that tests of middle difficulty would soon become too easy to provide adequate measures of achievement.
2. The tests are well adapted to the various levels of instruction. Steady increase in score is shown in the norms as the students progress from year to year of instruction and from lower to higher level form.
3. While for most languages the students in traditional programs do show some superiority in the reading and writing skills, the superiority is not so marked as critics of the audio-lingual method expected they might be. As a matter of fact, the superiority is so slight that no mention is made of it in the handbook that accompanies the tests.
4. Students with two years of high school instruction and one year of college instruction do about equally well on the tests, but students with four years of high school instruction do slightly better than students with two years of college instruction. This is no doubt attributable to a great degree to the fact that four years of high school instruction is, as a rule, offered in superior high schools and the best and most serious students of the language take the fourth year of instruction. The same level of ability and seriousness of purpose are not always present in college students completing a required second year of language study.

These, then, are the MLA Cooperative Foreign Language Tests — the tests with the "new look," which are apparently fulfilling most efficiently the "new purposes" that they were created to serve. They are currently being widely used in public secondary schools and in independent schools as measures of achievement, and they are being used in a large number of colleges and universities for placement purposes and end-of-course testing. They are also serving at all levels of foreign language instruction as models for teacher-made tests.

Their enthusiastic acceptance by the thousands of foreign language teachers who have used the MLA Cooperative Foreign Language Tests is indeed a fine tribute to the foreign language teaching profession as a whole — and especially to the professional organization which sponsored the development of the tests and the members of the profession who planned and created them.

---

*English teaching should pay more attention to the speaking, listening to, and creative uses of language for young people at all levels.* These were among the recommendations of a unique conference, the Dartmouth Seminar, which brought together 50 scholars and specialists in the teaching of English in Anglo-American countries for a month-long study. Participants generally agreed that if there is a "new English," it is to be found in reexamining and reinterpreting a child's experiences in language, not by introducing new content. Teachers should say less and children more in English classrooms, and there should be many opportunities for creativity, e.g., creative dramatics, imaginative writing, improvisation, and role playing. The Seminar criticized the rigidity of "grouping" or "streaming," as these "limit the linguistic environment in which boys and girls learn English." Also, present examination patterns direct the attention of both teachers and pupils "to aspects of English which are at best superficial and often misleading." The Seminar was financed by the Carnegie Corporation and cosponsored by the Modern Language Association, the National Council of Teachers of English, and Great Britain's National Association for the Teaching of English.

— *Education U.S.A.*, October 20, 1966